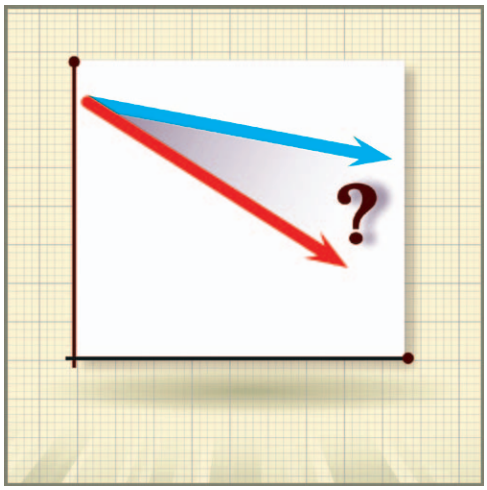
How Much Better Is Good Enough?

Patient-reported Outcomes, Minimal Clinically Important Differences, and Patient Acceptable Symptom States in Perioperative Research

Duminda N. Wijeysundera, M.D., Ph.D., Sindhu R. Johnson, M.D., Ph.D.

MORTALITY after surgery is increasingly rare in developed countries.[1] Nonetheless, many opportunities exist to improve perioperative care. For example, surgical patients still need postoperative comfort that involves minimization of postoperative pain, prevention of nausea, and facilitation of early mobilization. Many patients continue to experience major nonfatal complications that result in long-term disability and poorer quality of life. Given this context, the patient-reported outcome (PRO) is likely to take on an increasingly important role in perioperative research. For example, several available scales can now integrate and quantify the multiple dimensions constituting quality of recovery (QoR) after surgery.[2,3] Disability and health-related quality of life scales can assess the impact of postoperative morbidity on patients' ability to return to normal function and quality of life.[4,5] While PROs that have undergone rigorous psychometric evaluation can help better assess patients' perioperative experience, they also introduce new challenges. A critical issue is determining the minimal clinically important difference (MCID) or the smallest difference in PRO scores that "*patients perceive as beneficial.*"[6] Taking the example of the commonly used numeric rating scale of pain intensity, what reduction in pain scores represents a meaningful improvement in postoperative pain for a patient? This information can help physicians assess whether treatment effects in a randomized trial are clinically relevant and help researchers more accurately estimate sample sizes for any study that includes a PRO.[7] In this issue of Anesthesiology, Myles *et al.*[8] present a high-quality prospective cohort study that determined the MCID of three postoperative QoR scales, namely the QoR Score, QoR-15, and QoR-40. The investigators recruited 204 patients undergoing a broad range of surgical procedures at three Australian hospitals. The participants completed the three QoR scales at several postoperative time points, with 199 individuals undergoing at least 2 postoperative interviews. Based on several different analyses, the investigators found the MCID to be 0.9 for the QoR Score, 8.0 for the QoR-15, and 6.3 for the QoR-40. This well-conducted study is likely to form the methodologic foundation for many future studies evaluating QoR in surgical patients. Given its importance, and the paucity of similar studies in the anesthesiology literature, we will focus the remainder of this editorial on typical methods for estimating the MCID and future directions for such research in the perioperative arena.



***"[An increasingly important issue in perioperative research will be] determining the minimal clinically important difference ... or the smallest difference in [patient-reported outcome scores] that* 'patients perceive as beneficial.'"**

Methods for Estimating Clinically Important Changes in Patient-reported Outcomes

Myles *et al.* used four methods in tandem to calculate the MCID and averaged the results to derive a final overall

Image: A. Johnson.

Corresponding article on page 39.

Accepted for publication March 20, 2016. From the Li Ka Shing Knowledge Institute of St. Michael's Hospital, Toronto, Ontario, Canada (D.N.W.); Department of Anesthesia, Toronto General Hospital and University of Toronto, Toronto, Ontario, Canada (D.N.W.); Institute of Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada (D.N.W., S.R.J.); and Division of Rheumatology, Toronto Western Hospital, Mount Sinai Hospital and University of Toronto, Toronto, Ontario, Canada (D.N.W., S.R.J.).

 Anesthesiology 2016; 125:7-10

<zdoi;10.1097/ALN.0000000000001159>

estimate. Several commonly used approaches were employed, namely three *distribution-based* methods and an *anchor-based* method. This strategy of using several different estimation methods simultaneously is consistent with current recommendations.[9] Specifically, there is no single optimal method for estimating MCIDs, with each approach having strengths and limitations.[10] Distribution-based methods are used to calculate the MCID using the statistical distribution (*e.g.*, proportion of SD) and reliability (*e.g.*, SE of measurement) of the outcome measure. Their major advantage is convenience. Estimates can be readily calculated once the relevant scale has been implemented in a representative cohort. Conversely, distribution-based methods do not explicitly relate the MCID estimate to changes in patients' self-reported experience. Thus, these methods may not consistently identify the smallest difference in scores that "*patients perceive as beneficial.*"[11]

An alternative framework for estimating MCID involves comparing changes in PRO scores to a patient-reported "anchor," typically an additional single question asking patients to characterize any change in their overall health status. For example, during the follow-up visit, patients might be asked to rate any global improvement or worsening of their health-related quality of life using a Likert scale. The MCID is then estimated based on changes in PRO scores among patients who also reported a slight improvement *or* worsening of their global health status. The major advantage of the anchor-based approach is its explicit linkage of the MCID estimate to patients' self-reported experience. Conversely, it has important disadvantages. The MCID estimate can be imprecise, especially since these analyses only include individuals who report a slight improvement or worsening in their global health status. For example, in this current study, while 199 individuals underwent at least 2 postoperative interviews using QoR scales, only 51 patients contributed data to the anchor-based estimate of MCID. The anchor-based method is strongly influenced by patients' current state, and only weakly by their previous state.[7] Furthermore, patients may not accurately recall their baseline health status, especially in the immediate postoperative setting where many patients are exposed to sedative or opioid medications.

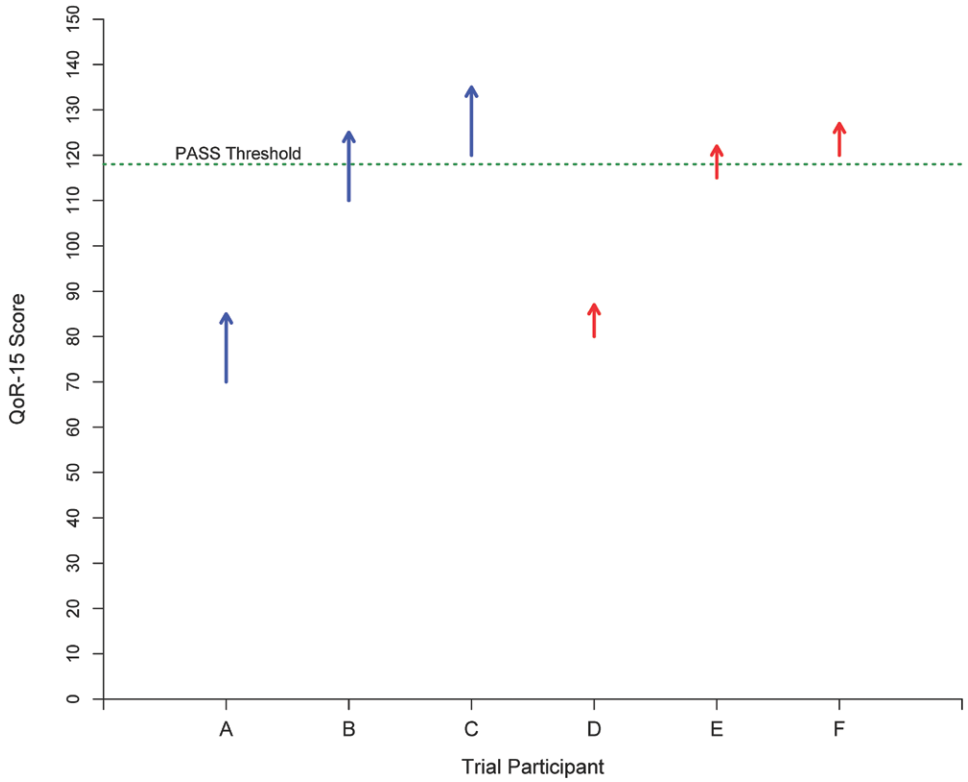


Fig. 1. Trial participants in a hypothetical randomized trial using a patient-reported outcome. The *arrows* depict the increase in quality of recovery (QoR), as measured by QoR-15 scores, of six participants in a hypothetical randomized trial. The QoR-15 score has a range of 0 to 150, with a minimal clinically important difference (MCID) of 8.0 and patient acceptable symptom state (PASS) of 118. Increases in QoR-15 scores are consistent with improved QoR. The *green dotted horizontal line* denotes the PASS, namely, the minimal absolute threshold in QoR-15 scores consistent with a good quality of postoperative recovery. Participants A, B, and C (denoted in *blue*) all experienced improvements that exceeded the MCID. Nonetheless, only participants B and C eventually attained a good overall QoR by exceeding the PASS threshold. Conversely, participants D, E, and F (denoted in *red*) did not experience clinically significant improvements in QoR-15 scores (*i.e.*, increases in QoR-15 scores that did not exceed the MCID). Nonetheless, participants E and F still eventually attained a good quality of postoperative recovery (*i.e.*, final scores exceeding PASS threshold). The concept for this figure was adapted from a presentation by Beaton[14] (http://www.immpact.org/static/meetings/Immpact4/background/beaton.pdf).

Feeling Better *versus* Feeling Good

Importantly, both distribution-based and anchor-based methods for calculating MCID focus on the *change* in patients' health status. For example, the MCID implicitly assumes that patients would consider a reduction in numeric rating scores for acute pain from eight points to six points to be equivalent to a reduction from five points to three points. This may not be the case. Satisfaction with pain control might also require that the pain score decreases below a specific threshold. This threshold has been termed the patient acceptable symptom state (PASS) or the PRO value beyond which patients consider themselves well.[12] As described by Tubach *et al.*,[13] the "*MCID deals with the concept of improvement (feeling better)*," while PASS addresses the concept of "*wellbeing or remission of symptoms (feeling good).*" Since both constructs are highly relevant to patients, Myles *et al.* conducted secondary analyses that showed attainment of good postoperative recovery to be consistent with a QoR Score of 16, QoR-15 of 118, or QoR-40 of 180.

Moving Forward

As PROs take on an increasingly important role in evaluating perioperative outcomes, how should related methodologic research move forward? *First*, once valid estimates of MCID and PASS are known, this information should be used to help readers interpret the clinical relevance of any randomized trial that used a PRO (fig. 1).[14] For example, future trials that use QoR-15 as an outcome should report not only aggregate changes in QoR-15 scores (*e.g.*, mean and SD) but also the proportions of patients who attained improvements exceeding the MCID (*i.e.*, 8.0 reduction or greater) and who attained the PASS (*i.e.*, score of 118 or greater). *Second*, more high-quality research is needed to estimate MCIDs of PROs, including validation studies that reestimate MCIDs in other settings. The MCIDs estimated by Myles *et al.* are representative of mostly English-speaking (83%) patients at three Australian hospitals where 89% of individuals experienced a good postoperative recovery. Since there is likely no single MCID that applies to all populations, future studies must determine whether these estimates can be extrapolated to different populations and settings. *Third*, the evidence base surrounding PROs for use in the perioperative setting will need to be summarized at regular intervals to identify the most psychometrically valid options for measuring specific constructs (*e.g.*, QoR, disability, and burden of complications), along with recommended MCID and PASS estimates. These reviews are best linked to ongoing initiatives to standardize endpoints in perioperative research,[15] and should be informed by similar programs in other areas of medicine such as chronic pain and rheumatology.[16,17] This long-term effort to better measure and characterize the self-reported experience of surgical patients will help ensure that our patients will live *more often*, live *better*, and live *well* after surgery.

Research Support

Supported by a New Investigator Award from the Canadian Institutes of Health Research, Ottawa, Ontario, Canada (to Dr. Wijeysundera); a Merit Award from the Department of Anesthesia at the University of Toronto, Toronto, Ontario, Canada (to Dr. Wijeysundera); and a Clinician Scientist Award from the Canadian Institutes of Health Research, Ottawa, Ontario, Canada (to Dr. Johnson).

Competing Interests

The authors are not supported by, nor maintain any financial interest in, any commercial activity that may be associated with the topic of this article.

Correspondence

Address correspondence to Dr. Wijeysundera: d.wijeysundera@utoronto.ca

References

1. Bainbridge D, Martin J, Arango M, Cheng D; Evidence-based Peri-operative Clinical Outcomes Research (EPiCOR) Group: Perioperative and anaesthetic-related mortality in developed and developing countries: A systematic review and meta-analysis. Lancet 2012; 380:1075–81
2. Wong J, Tong D, De Silva Y, Abrishami A, Chung F: Development of the functional recovery index for ambulatory surgery and anesthesia. Anesthesiology 2009; 110:596–602
3. Myles PS, Hunt JO, Nightingale CE, Fletcher H, Beh T, Tanil D, Nagy A, Rubinstein A, Ponsford JL: Development and psychometric testing of a quality of recovery score after general anesthesia and surgery in adults. Anesth Analg 1999; 88:83–90
4. Shulman MA, Myles PS, Chan MT, McIlroy DR, Wallace S, Ponsford J: Measurement of disability-free survival after surgery. Anesthesiology 2015; 122:524–36
5. Mangione CM, Goldman L, Orav EJ, Marcantonio ER, Pedan A, Ludwig LE, Donaldson MC, Sugarbaker DJ, Poss R, Lee TH: Health-related quality of life after elective surgery: Measurement of longitudinal changes. J Gen Intern Med 1997; 12:686–97
6. Jaeschke R, Singer J, Guyatt GH: Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989; 10:407–15
7. Streiner DL, Norman GR, Cairney J: Health Measurement Scales: A Practical Guide to Their Development and Use. New York: Oxford University Press, 2014
8. Myles PS, Myles DB, Galagher W, Chew W, MacDonald N, Dennis A: The minimal clinically important difference for three quality of recovery scales. Anesthesiology 2016; 125:39–45
9. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR; Clinical Significance Consensus Meeting Group: Methods to explain the clinical significance of health status measures. Mayo Clin Proc 2002; 77:371–83
10. King MT: A point of minimal important difference (MID): A critique of terminology and methods. Expert Rev Pharmacoecon Outcomes Res 2011; 11:171–84
11. Farivar SS, Liu H, Hays RD: Half standard deviation estimate of the minimally important difference in HRQOL scores? Expert Rev Pharmacoecon Outcomes Res 2004; 4:515–23

12. Dougados M: It's good to feel better but it's better to feel good. J Rheumatol 2005; 32:1–2
13. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, Bombardier C, Felson D, Hochberg M, van der Heijde D, Dougados M: Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: The patient acceptable symptom state. Ann Rheum Dis 2005; 64:34–7
14. Beaton DE. Clinical Importance: The OMERACT Perspective. 2004. Available at: http://www.immpact.org/static/meetings/Immpact4/background/beaton.pdf. Accessed March 25, 2016
15. Myles PS, Grocott MP, Boney O, Moonesinghe SR, on behalf of the COMPAC-StEP Group. Standardizing end points in perioperative trials: Towards a core and extended outcome set. Br J Anaesth 2016; 116:586–9
16. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L: OMERACT: An international initiative to improve outcome measurement in rheumatology. Trials 2007; 8:38
17. Dworkin RH, Turk DC, McDermott MP, Peirce-Sandner S, Burke LB, Cowan P, Farrar JT, Hertz S, Raja SN, Rappaport BA, Rauschkolb C, Sampaio C: Interpreting the clinical importance of group differences in chronic pain clinical trials: IMMPACT recommendations. Pain 2009; 146:238–44

ANESTHESIOLOGY REFLECTIONS FROM THE WOOD LIBRARY-MUSEUM

Galen's Four Temperaments in 1893 at Chicago's Post-Graduate School of Anaesthesia

Born over 500 years after the death of Hippocrates, Galen of Pergamon (129 CE to c.216 CE), a Greek physician to several Roman emperors, reoriented medicine from the Hippocratic humoral system toward a Galenic temperamental one. From animal sacrifices to the gods, pre-Classical Greeks had observed that the topmost layer (serum) of freshly settled sacrificial blood was actually yellow. This **leading** layer likely sparked concepts of elemental fire, Hippocratic "yellow bile," and the Galenic "choleric" temperament of **leading**, goal-oriented motivators. The second buff layer (leukocytes) reflected elemental **water**, Hippocratic "phlegm," and the Galenic "phlegmatic" temperament of **water**-like, calming mediators. The third **air**-reddened layer (erythrocytes) inspired concepts of elemental **air**, Hippocratic "blood," and the Galenic "sanguine" temperament of garrulous, "**air**-filling" talkers. The final, reddish-black **bottom** layer (clot) congealed elemental earth, Hippocratic "black bile," and the Galenic "melancholic" temperament of depressed, **bottom**-line, analytical perfectionists. In 1893 at Chicago's Post-Graduate School of Anaesthesia, professors taught that "Physiognomy, Phrenology and Temperament" provided clues to each patient's unique mixture of Galen's choleric, phlegmatic, sanguine, and melancholic temperaments.

George S. Bause, M.D., M.P.H., Honorary Curator, ASA's Wood Library-Museum of Anesthesiology, Schaumburg, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYC@aol.com.